DOCUMENT RESUME

ED 395 264 CG 027 091

AUTHOR Pajares, Frank
TITLE Assessing Self-Efficacy Beliefs and Academic Outcomes: The Case for Specificity and Correspondence.
PUB DATE 96
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Academic Aspiration; Classroom Research; *Educational Research; Elementary Secondary Education; Evaluation; Expectation; Individual Power; Learning Theories; Secondary Education; *Self Efficacy; Self Esteem; Self Motivation

ABSTRACT

Guidelines regarding self-efficacy assessment are highlighted in the first section of this paper. In the second section, the issue of specificity versus generality of measurement is clarified. And last, preliminary results of a study of eighth graders (n=172) are presented which demonstrate that: (1) the optimal level of specificity of any efficacy assessment depends on the complexity of the performance criteria with which it is compared; and (2) judgments of competence need not be so microscopically operationalized that the assessment loses all sense of practical utility. When a criterial task of interest is relatively broad, such as term grades, self-efficacy judgments can be tailored to these levels and still remain highly predictive. Various forms of self-referent thought measured at various levels of specificity can also prove useful outside the research arena as diagnostic and assessment tools--they can provide teachers and counselors with information regarding students' dispositions, and results may be useful in helping to understand affective influences on performances that do not easily lend themselves to microanalytic analysis. Table 1 presents Bandura's guidelines (1986) regarding the specificity and correspondence of self-efficacy and performance assessment. Table 2 presents sample self-efficacy items. A figure presents correlations between different levels of self-efficacy assessment and differing students performance outcomes. (TS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 395 264

Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence

Frank Pajares
Emory University

Please address correspondence to:

Frank Pajares
Educational Studies
Emory University
Atlanta, GA 30322

Tel: 404-727-1775
Fax: 404-727-2799
Email: mpajare@emory.edu

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY:
F. PAJARES
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at a symposium chaired by B. J. Zimmerman, *Measuring and mismeasuring self-efficacy: Dimensions, problems, and misconceptions*. Annual meeting of the American Educational Research Association, New York, 1996.

CG027091

BEST COPY AVAILABLE

Assessing self-efficacy beliefs and academic outcomes:
The case for specificity and correspondence

Bandura (1986, in press) has cautioned researchers attempting to predict academic outcomes from students' self-efficacy beliefs that, to increase accuracy of prediction, they would be well advised to follow theoretical guidelines regarding specificity of self-efficacy assessment and correspondence with criterial tasks. This caution has often gone unheeded in educational research, resulting in self-efficacy assessments that reflect global or generalized self-perceptions of competence and that bear slight resemblance to the criterial task with which they are compared. Often, no criterial task is identified, as researchers aim to discover simply the nature of the interplay between motivational variables in the absence of performance attainments. In still other studies, judgments of "confidence" that bear passing resemblance to self-efficacy judgments are used instead of more appropriate particularized measures. The result is often confounded relationships and ambiguous findings that obfuscate the potential contribution of self-efficacy beliefs to the understanding of academic performances.

I hope to accomplish three objectives during this presentation. First, I want to briefly highlight some of the guidelines provided by Bandura (in press) regarding self-efficacy assessment. Second, I want to clarify the issue of specificity versus generality of measurement. And last, I want to report on preliminary results of a study demonstrating that, when efficacy beliefs assessed do not reflect with specificity the academic tasks with which they are compared, their predictive value is diminished. Conversely, prediction of academic outcomes is enhanced as self-efficacy and performance more closely correspond.

First, let me highlight what I believe to be particularly useful guidelines presented by Bandura (in press) regarding the specificity and correspondence of self-efficacy and performance

assessment **[Overhead 1, Table 1]**. The broadest, most general, self-efficacy assessments would consist of an omnibus-type instrument that attempts to measure a general sense of "confidence." Such **omnibus measures** create problems of predictive relevance and are obscure about just what is being assessed. Omnibus tests that aim to assess general self-efficacy, for example, provide global scores that decontextualize the self-efficacy/behavior correspondence and transform self-efficacy into a generalized personality trait rather than the context-specific judgment Bandura (1986, in press) suggests it is. After all, generalized self-efficacy instruments basically assess "people's general belief that they can make things happen without specifying what [these things] are" (Bandura, in press). Even **domain-specific omnibus measures** are problematic if composite multiscale scores drawn from differing subsections of the domain are used. It is not altogether easy, for example, to see what value composite scores provided by multiple-scale instruments such as the often used Mathematics Self-Efficacy Scale may have, especially if one wishes to predict relatively discrete mathematics outcomes. The same might be said of a scale that would regard "writing self-efficacy" as the combined score of one subscale assessing a student's confidence to accomplish writing tasks of varying difficulty and another subscale assessing confidence to perform various composition skills.

Various researchers have assessed **general academic self-perceptions of competence**. The basic problem with such assessments is that students must generate these judgments without a clear academic activity or task in mind. As a result, they generate the judgments by in some fashion mentally aggregating related perceptions that they hope will be related to imagined tasks. Domain-specific assessments, such as asking students to provide their confidence to learn mathematics or writing, are more explanatory and predictive than omnibus measures and

preferable to general academic judgments, but they are inferior to task-specific judgments because the subdomains differ markedly in the skills required.

Academic domain-specific assessments of self-efficacy are especially common in educational research in part because the criterial outcome tasks such as semester grades or achievement test results that are often used do not lend themselves well to particularized self-efficacy assessment. The typical strategy of researchers in this regard is to use multiple items to restate different facets (or even similar facets differently phrased) of the same academic subject. For example, it is not unusual for a mathematics self-efficacy scale to be populated with items such as "I am confident about my ability to do the work in this class;" "I am certain I can understand the math presented in this class;" and "I am confident I can perform as well or better than others in this class." Although high internal consistence can be counted on, such an assessment primarily provides a redundant measure of the general domain of mathematics.

Reasonably **precise judgments of capability** matched to a specific outcome afford the greatest prediction and offer the best explanations of performance outcomes, for these are typically the sorts of judgments that individuals use when confronted with behavioral tasks (Bandura, 1986). To this end, if the purpose of an study is to achieve explanatory and predictive power, **self-efficacy judgments should be consistent with and tailored to the domain of functioning** and/or task under investigation. This is especially critical in studies that attempt to establish causal relations between beliefs and outcomes. All this is to say that **capabilities assessed and capabilities tested should be the same** capabilities. When these guidelines regarding correspondence between belief and outcome are not followed, the resulting loss of predictive power is ensured and the influence of self-efficacy is minimized. Some researchers, for example, have operationalized mathematics self-efficacy as students' judgments of their

capabilities to solve math *problems*, perform math-related *tasks*, and succeed in math-related *courses*--the three subscales of the Mathematics Self-Efficacy Scale--although these judgments of math capabilities are substantively and conceptually quite different. If, as Bandura (1986) argued, self-efficacy assessment should be consistent with the criterial task to be useful and predictive, what criterial task is consistent with a composite score that comprises judgments of confidence to succeed in mathematics courses as diverse as geometry and accounting, complete math-related tasks as disparate as filling out an income tax form and figuring out how much material to buy so as to make curtains, and solve algebra and geometry problems of varying difficulty?

When these differing judgments of mathematics capability are compared with differing math-related outcomes--ability to solve the problems on which self-efficacy is assessed and math-relatedness of academic majors--results confirm that Bandura's (1986) guidelines regarding the match of self-efficacy and performance assessment are well founded (see Pajares & Miller, 1995). Students' confidence to solve mathematics problems is a more powerful predictor of their ability to solve those problems than is their confidence to perform math-related tasks or their confidence to earn high marks in math-related courses. Similarly, their confidence to succeed in such courses is more predictive of their choice of majors that required them to take many of the math-related courses on which they express that confidence. One might also question the practical utility of administering a 52-item instrument when greater prediction may be had from a shorter instrument more closely matching the performance task.

Let me also note that the **skills required to accomplish the performance attainments that form the outcome assessment should be clear to the participant**. When students do not know with any degree of accuracy what it is they are expected to do, the judgments on which

they will base their capability to do it will be nebulous at best. When criterial tasks are unclear, what little prediction is obtained is primarily due to the similarity of tasks and the powers of generalizability individuals can bring to bear on their self-perceptions. In these cases, a researcher can do little more than predict from a domain-specific measure that taps common demands.

Despite the optimal benefits that result from particularized efficacy and performance assessments, many academic outcomes of interest are not as particularized as, say, one's capability to solve specific mathematics problems, the typical level of specificity at which self-efficacy judgments are most predictive of academic performances. Lent and Hackett (1987) rightly observed that specificity and precision are often purchased at the expense of external validity and practical relevance. However, to be both practically useful and predictive, the **level of specificity of an efficacy assessment should depend on the complexity of the performance criteria with which it is compared**. Judgments of competence need not be so microscopically operationalized that the assessment loses all sense of practical utility (see Table 2 for sample self-efficacy items). If the criterial task should be a particularly relevant one such as choice of intention to enroll in math-related courses, self-efficacy judgments can be tailored to this level and still remain highly predictive. Lent, Lopez, and Bieschke (1993) showed how this can be accomplished when they compared students' confidence to succeed in math-related courses with three career-related outcomes--intention to take the courses listed on the instrument, grades obtained in math-related courses that students took during the subsequent term, and interest in the math courses listed on the instrument. Self-efficacy beliefs were predictive on each account, and such judgments offer information not afforded by broader judgments of competence.

Moreover, researchers have demonstrated that self-efficacy perceptions are also good predictors of reasonably generalized performances such as obtained grades (Bandura, 1993; Zimmerman, Bandura, & Martinez-Pons, 1991) or choice of academic majors (Hackett & Betz, 1989), and it bears repeating that the optimal level of specificity of any efficacy assessment should ultimately depend on the complexity of the performance criteria with which it is compared (see Lent & Hackett, 1987).

Let me close by briefly reporting on results of a study demonstrating that prediction of academic outcomes is enhanced as self-efficacy and performance more closely correspond **[Overhead 2, Figure 1]**. Colleagues and I are currently engaged in a study that in part attempts to investigate the relationship between mathematics self-efficacy beliefs assessed at three levels of specificity and their corresponding outcomes. To this end, we ask 8th grade students to report their confidence (a) to correctly solve each of 20 problems on a high-stakes mathematics exam, (b) in the various letter grades they will earn on the exam after completing the instructional unit but without seeing the specific exam questions, and (c) in the class grade they will receive for math at end of term. The self-efficacy assessments and the exam take place during the first week of a 9-week term. You will, I know, permit me to paint the canvas of this preliminary report with broad strokes as data input and collection is still ongoing.

Preliminary findings from our first assessment reveal that, when belief assessed and criterial task are matched, prediction is enhanced. Moreover, the strongest prediction is obtained from the most particularized match. In this study, when confidence to solve specific problems is compared to the number of problems solved (.57), the relationship is stronger than, say, when confidence in the overall mathematics test grade is compared to the grade on the test itself (.42). Also, note that the relationship between matched beliefs and outcomes is generally stronger than

between unmatched variables *even when the efficacy assessment is particularized*. For example greater prediction of term grade is available from self-efficacy beliefs about term grades (.45) than from self-efficacy to solve problems (.34). As Bandura (in press) has outlined, the issue is one of specificity *and* correspondence. The Williams T_2 statistic was used to determine that the correlations were significantly different

These findings demonstrate that the optimal level of specificity of any efficacy assessment depends on the complexity of the performance criteria with which it is compared, and that judgments of competence need not be so microscopically operationalized that the assessment loses all sense of practical utility. When a criterial task of interest is relatively broad, such as term grades, self-efficacy judgments can be tailored to these levels and still remain highly predictive. Various forms of self-referent thought measured at various levels of specificity can also prove useful outside the research arena as diagnostic and assessment tools--they can provide teachers and counselors with information regarding students' dispositions, and results may be useful in helping to understand affective influences on performances that do not easily lend themselves to microanalytic analysis.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Bandura, a. (in press). *Self-efficacy: The exercise of control.* New York: Freeman.

Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education, 20*, 261-273.

Lent, R. W., & Hackett, G. (1987). Career self-efficacy: Empirical status and future directions. *Journal of Vocational Behavior, 30*, 347-382.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1993). Predicting mathematics-related choice and success behaviors: Test of an expanded social cognitive model. *Journal of Vocational Behavior, 42*, 223-236.

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics outcomes: The need for specificity of assessment. *Journal of Counseling Psychology, 42*, 190-198.

Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society, Series B, 21*, 396-399.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663-676.

Table 1 **ON SPECIFICITY OF JUDGMENT**

Omnibus Measures are Problematic - Omnibus measures create problems of predictive relevance as well as obscurity about what is being assessed.

General Academic Judgments are Problematic - When an academic domain is unspecified, students have to come up with a single judgment of their capabilities from the scholastic activity that happens to spring to mind at the moment, guess about the scholastic activity the assessor had in mind, or engage in subjective weighting and aggregation across diverse subject matters.

Domain-specific Judgments are Preferable to Generalized Judgments but Inferior to Task-specific Judgments - A self-efficacy measure cast in terms of a general academic domain is more explanatory and predictive than an omnibus measure but still deficient because scientific, mathematics, linguistic, literary and artistic academic subdomains differ markedly in the types of competencies they require.

Academic Domain-specific Assessments Using Multiple Items Restating Different Facets of the Same Domain are Problematic - Items that measure perceived self-efficacy for different facets of the same domain cannot be treated as equivalent. Restricting items to those that correlate highly with each other may produce a self-efficacy scale that measures redundantly only a segment of the intended domain, and a narrow range of the segment at that.

Particularized Judgments are the Most Predictive Assessments - Efficacy beliefs should be measured in terms of particularized judgments of capability that may vary across realms of activity, different levels of task demands within a given activity domain, and under different situational circumstances. Major progress in understanding how personal factors operate in causal structures requires explicit measurement of the particular personal determinants that are germane to given spheres of functioning.

Particularized Level of Generality Should be Determined by the Nature of the Criterial Task - Self-efficacy is commonly misconstrued as concerned solely with "specific behaviors in specific situations." Domain particularity need not mean behavioral specificity. Efficacy beliefs are multifaceted and contextual, but the level of generality of the efficacy items within a given domain of functioning varies depending on the degree of situational resemblance and foreseeability of task demands. *Bur regardless of the level of generality, in no case are the efficacy items disembodied from contexts and level of task demands.*

Adapted from Albert Bandura (in press)
Self-Efficacy: The Exercise of Control

Table 1 (continued)

ON CORRESPONDENCE BETWEEN JUDGMENT AND OUTCOME

Self-efficacy Should be Consistent with and Tailored to the Performance Outcome With Which it is Compared - To achieve explanatory and predictive power, efficacy measures must be tailored to domains of functioning and represent gradations of task demands within those domains. Causal processes are best clarified by microanalytic measures of efficacy beliefs tailored to the domain of functioning being explored.

Capabilities Assessed and Tested Should Match - The structure of the relation between efficacy beliefs and action requires that both events tap similar capabilities.

Skills Required to Accomplish the Criterial Task Should be Clear to the Participant - Judgments of self-efficacy require knowledge of task demands. If one does not know what demands must be fulfilled in a given endeavor, one cannot accurately judge whether one has the requisite capabilities to do so.

When Criterial Task is Unknown, Self-efficacy Assessment with Generalized Measure Results in Loss of Predictive Power - When the situations people are likely to encounter are not fully known, one would do better predicting from perceived efficacy for common than for unusual situational demands. It should be noted, however, that efficacy predictiveness for common conditions is gained at the loss of some predictiveness for conditions within the same general context that have fewer common features.

Adapted from Albert Bandura (in press)
Self-Efficacy: The Exercise of Control

Table 1 (continued)

ASSESSING SELF-EFFICACY AND CRITERIAL TASK

High-stakes Assessments are Preferable - Individuals may possess the requisite skills and a strong sense of efficacy that they can execute them well, but choose not to perform the activities because they have no incentive to do so. Situations in which misjudgments of capabilities carry no consequences provide little incentive for serious appraisal of personal efficacy

Assessments Should Take Place in Close Temporal Proximity - The existing relation between efficacy beliefs and action is most accurately revealed when they are measured in close temporal proximity. The closer in time the better the test for causation.

Self-efficacy Assessment Should Encompass the Full Range of Criterial Task - Self-efficacy theory is often tested not on a completely unfounded factor, but on one that exerts only partial influence over the behavior of interest.

Specificity and Correspondence - Causal processes are best clarified by microanalytic measures of self-efficacy beliefs that are tailored to the domain of functioning being explored.

Items Should be Phrased in Terms of *Can* rather than *Will* - The items are phrased in terms of *can* rather than *will*. Can is a judgment of capability; will is a statement of intention.

Items Should be Phrased in Terms of *Can Now* rather than *Can Later* - People are asked to judge their operative capabilities as of now not their potential capabilities or their expected future capabilities.

Adapted from Albert Bandura (in press)
Self-Efficacy: The Exercise of Control

Table 1 (continued) **SOURCES OF DISCORDANCE BETWEEN EFFICACY JUDGMENT AND ACTION**

Disincentives and Performance Constraints - Individuals may possess the requisite skills and a strong sense of efficacy that they can execute them well, but choose not to perform the activities because they have no incentive to do so.

Temporal Disparities - The existing relation between efficacy beliefs and action is most accurately revealed when they are measured in close temporal proximity. The closer in time the better the test for causation.

Mismatch or Partial Match of Assessed Capabilities - The structure of the relation between efficacy beliefs and action requires that both events tap similar capabilities.

Limited Scope of Self-efficacy Assessment - Self-efficacy theory is often tested not on a completely unfounded factor, but on one that exerts only partial influence over the behavior of interest.

Faulty Assessment of Self-efficacy or Performance - Causal processes are best clarified by microanalytic measures of self-efficacy beliefs that are tailored to the domain of functioning being explored.

Ambiguity of Task Demands - Judgment of self-efficacy requires knowledge of task demands. If one does not know what demands must be fulfilled in a given endeavor, one cannot accurately judge whether one has the requisite capabilities to do so.

Indefinite Aims and Deficient Performance Information - Efficacy beliefs cannot operate as a regulative influence in an informational vacuum. The problem of performance ambiguity arises when important aspects of one's performance are not personally observable.

Consequences of Misjudgment - Situations in which misjudgments of capabilities carry no consequences provide little incentive for serious appraisal of personal efficacy.

Faulty Self-knowledge - In new undertakings, people have a limited basis on which to assess the adequacy of their self-appraisals. Self-efficacy can also be misjudged when personal factors distort self-appraisal processes. Distortions in memory of efficacy-relevant experiences and the circumstances under which they occurred will produce faulty self-appraisals.

Adapted from Albert Bandura (in press)
Self-Efficacy: The Exercise of Control

Table 2 **Items Used to Measure Various Self-Efficacy Constructs**

Source	Sample question or direction	Answer options
Teaching efficacy (Bandura, 1993)	How much can you [followed by various teaching-related tasks--e.g., "influence the decisions that are made in your school"]	(1) *nothing* to (9) *a great deal*--in intervals of 1
Mathematics problem-solving self-efficacy (Pajares & Miller, 1994)	How confident are you that you could give the correct answer to the following problem without using a calculator? [followed by 20 algebra or geometry problems--e.g., "Simplify: -6[x + (-7y)] + (-5)(3x - y)"]	(1) *no confidence* to (6) *complete confidence*--in intervals of 1
Self-efficacy for self-regulated learning (Bandura, 1989)	How well can you [followed by 11 self-regulatory tasks--e.g., "finish homework assignment by deadlines"]	(1) *not well at all* to (7) *very well*--in intervals of 1
Self-efficacy for writing skills (Shell et al., 1989)	How confident are you that you can perform each of the following writing skills? [8 skills presented--e.g, "correctly spell all words in a one page passage"]	Scale of 0 to 100 - student writes the specific number
Mathematics courses self-efficacy (Betz & Hackett, 1983)	How much confidence do you have that you could complete the following course with a final grade of B or better? [courses presented--"e.g., algebra]	(0) *no confidence* to (9) *complete confidence*--in intervals of 1
Collective efficacy (Bandura, 1993)	Please indicate your confidence that you can attain the following *grade level gains* with the students in your class this year. [gains in 2-month increments presented]	(0) *no confidence at all* to (10) *certain I can do*--in intervals of 1
Self-efficacy for division problems (Schunk, 1981)	[Division problem shown for 2 seconds] Circle the number on the line that matches how sure you are that you could work problems like those shown and get the right answers.	(10) *not sure* to (100) *real sure* in--intervals of 10
Self-efficacy for reading tasks (Shell et al., 1995)	How confident are you that you can perform each of the following reading tasks? [18 tasks presented--e.g., "read a letter from a friend"]	(1) *I'm sure I can* to (5) *I'm sure I can't*--in intervals of 1
Self-efficacy for academic achievement (Bandura, 1989)	How well can you [followed by 9 academic domains--e.g., "learn general mathematics, learn reading and writing language skills"]	(0) *not well at all* to (7) *very well*-- in intervals of 1

Bandura, A. (1989). *Multidimenstional scales of perceived self-efficacy*. Unpublished test. Stanford University, Stanford, CA.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*, 329-345.

Pajares, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis, *Journal of Educational Psychology, 73*, 93-105.

Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology, 87*, 386-398.

Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*, 91-100.

BEST COPY AVAILABLE

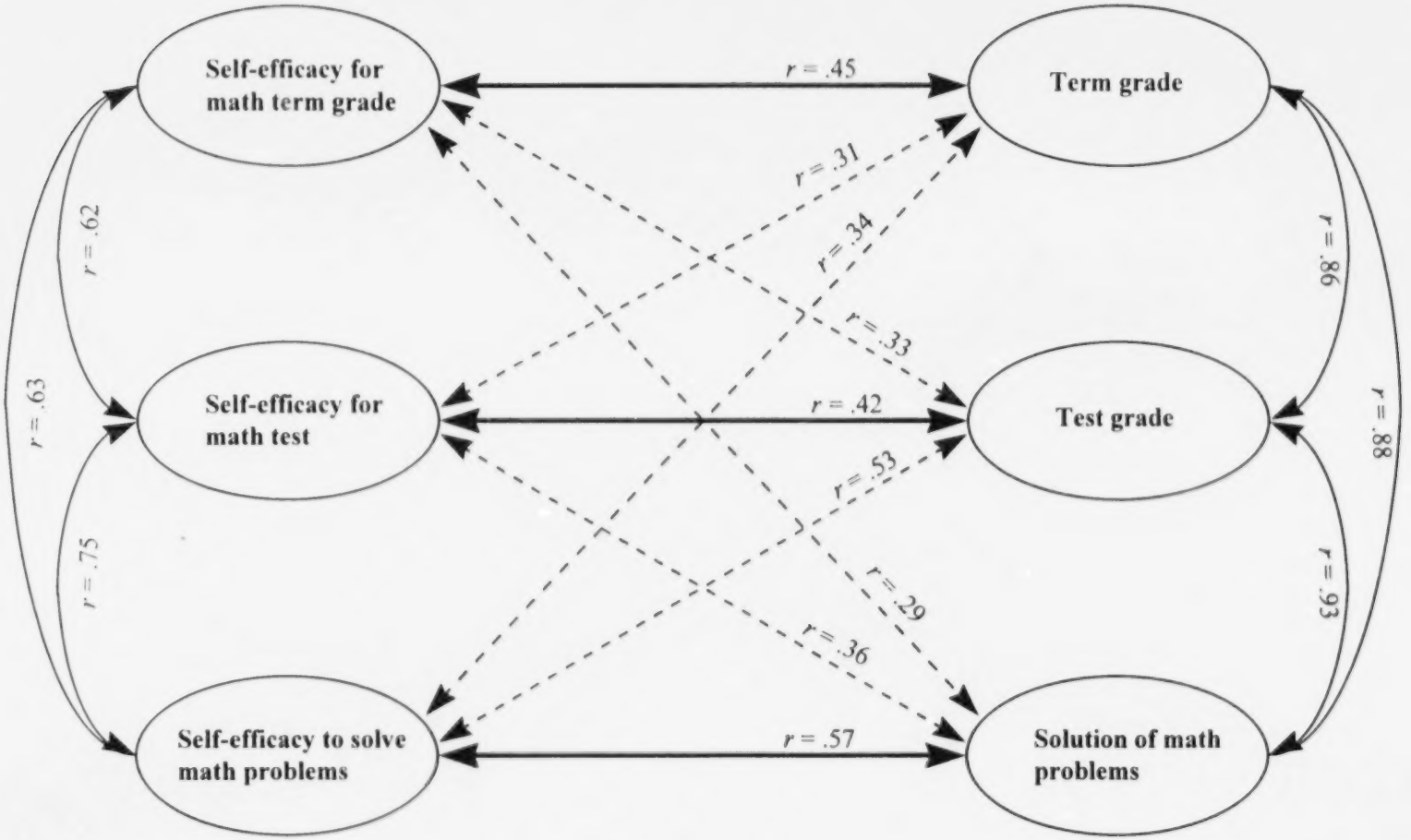

Figure 1 - Zero-order correlations between different levels of self-efficacy assessment and differing mathematics outcomes (N = 172).